High-Quality Draft Genome Sequences of 28 Enterococcus sp. Isolates

JOURNAL OF BACTERIOLOGY, May 2010, p. 2469–2470 Vol. 192, No. 9
0021-9193/10/$12.00 doi:10.1128/JB.00153-10

High-Quality Draft Genome Sequences of 28 *Enterococcus* sp. Isolates▽

Kelli L. Palmer,[1,2,3] Karen Carniol,[1,2]† Janet M. Manson,[1,2]‡ David Heiman,[4] Terry Shea,[4] Sarah Young,[4] Qiandong Zeng,[4] Dirk Gevers,[4] Michael Feldgarden,[4] Bruce Birren,[4] and Michael S. Gilmore[1,2,3]*

Departments of Ophthalmology, and Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts[1]; Schepens Eye Research Institute, Boston, Massachusetts[2]; The Harvard Microbial Sciences Initiative, Cambridge, Massachusetts[3]; and The Broad Institute, Cambridge, Massachusetts[4]

Received 12 February 2010/Accepted 23 February 2010

The enterococci are low-GC Gram-positive bacteria that have emerged as leading causes of hospital-acquired infection. They are also commensals of the gastrointestinal tract of healthy humans and most other animals with gastrointestinal flora and are important for food fermentations. Here we report the availability of draft genome sequences for 28 enterococcal strains of diverse origin, including the species *Enterococcus faecalis*, *E. faecium*, *E. casseliflavus*, and *E. gallinarum*.

The enterococci are ecologically diverse, Gram-positive lactic acid bacteria that are found in the gastrointestinal consortia of humans, other mammals, reptiles, amphibians, birds, and insects and are utilized in production of fermented foods and probiotics (1, 2, 16). However, they have also emerged as leading causes of hospital-acquired infection at extraintestinal sites, including the heart, urinary tract, and surgical site wounds (9, 10). These infections have become a leading health care concern because of increasing antibiotic resistance mediated by mobile genetic elements (12). The ability of enterococci to colonize extraintestinal sites and cause infection involves both core and variable genetic traits (3, 15) and is not fully understood.

Enterococci exhibit diversity in metabolic capabilities (such as the ability of *Enterococcus faecalis*, but not *E. faecium*, to respire [11]) and in clinically relevant phenotypes such as toxin production (12) and biofilm formation (13). Enterococci can also be motile or nonmotile, and pigmented or nonpigmented (4), yet little is known about the mechanism and significance of these traits. Insights into basic enterococcal physiology are limited by a paucity of genome data. The genome sequences of the vancomycin-resistant bloodstream isolate *E. faecalis* V583 (14), the human oral isolate *E. faecalis* OG1RF (3), and the commercial probiotic strain *E. faecalis* Symbioflor (6) have been determined, and a draft genome sequence for *E. faecium* DO has been available in GenBank since 2002. Here we announce the availability of draft genome sequences of 28 additional enterococcal strains, including 16 *E. faecalis* strains, 8 *E. faecium* strains, 3 *E. casseliflavus* strains, and 1 *E. gallinarum* strain. In addition to expanding the *Enterococcus* nucleotide sequence in GenBank by approximately an order of magnitude, this is the first report of genome sequencing of the motile *E. gallinarum* strain and the motile, pigmented *E. casseliflavus* strains.

Most of the enterococcal strains utilized for our genome sequencing project are of clinical origin. The remaining strains are commensal isolates or were obtained from animals or insects. *E. faecalis* strains selected for this project were previously described in a study of *E. faecalis* strain diversity (12). To maximize the information yield from genome sequencing, we selected 16 *E. faecalis* strains that represent the deepest phylogenetic nodes in the *E. faecalis* multilocus sequence typing (MLST) dendrogram from the 106 strains previously examined (12). These strains include commensal isolates such as the human fecal isolate E1Sol, derived from a population with little exposure to antibiotics (8), and the fruit fly isolate Fly1 (5), as well as multiple antibiotic-resistant clinical isolates such as HIP11704, the strain implicated in the first documented interspecies transfer of vancomycin resistance genes to methicillin-resistant *Staphylococcus aureus* (7, 17). Six of the eight *E. faecium* strains and the *E. casseliflavus* and *E. gallinarum* strains utilized for genome sequencing were obtained from a clinical isolate repository (Eurofins Medinet) and were selected based on diversity in geographic isolation and antibiotic resistance phenotype. The remaining two *E. faecium* strains, Com12 and Com15, were isolated from the feces of healthy human volunteers. A full description of these strains will be included in a future report with the results of a comparative enterococcal genome analysis.

Genomes were sequenced to at least 15-fold coverage using 454 FLX pyrosequencing (Roche) with DNA fragment libraries and 3-kb paired-end reads according to the manufacturer's recommendations. Genomes were assembled using Newbler. Before assembly, quality of the 454 sequencing data was analyzed, and suspect libraries were removed. The runAssembly script was then used to assemble reads into contigs and scaffolds using the parameters −ar −rip −g. Each assembly was scanned for 16S rRNA sequence using the Ribosomal Database Project (RDP) to confirm the presence of one appropriately classified 16S rRNA sequence. Final assemblies were BLASTed to the NCBI nonredundant (NR) database,

* Corresponding author. Mailing address: Schepens Eye Research Institute, 20 Staniford Street, Boston, MA 02114. Phone: (617) 912-7422. Fax: (617) 912-0115. E-mail: michael.gilmore@schepens.harvard.edu.

† Present address: Cell Press, 600 Technology Square, Cambridge, MA 02139.

‡ Present address: Department of Biology, Northeastern University, 360 Huntington Avenue, Boston, MA 02115.

▽ Published ahead of print on 5 March 2010.

UniVecCore, and a mitochondrial database to remove any contaminating sequence. Annotation of assemblies was performed using a combination of *ab initio* and evidence-based approaches as described at the Broad Institute *Enterococcus* group database (http://www.broadinstitute.org/annotation/genome/enterococcus_faecalis/GeneFinding.html). A summary of gene finding data for each predicted locus can also be viewed at the Broad Institute *Enterococcus* group database.

Nucleotide sequence accession numbers. GenBank accession numbers (in parentheses) for genomes within this project are as follows: *E. casseliflavus* EC10 (ACAL00000000), EC20 (ACAO00000000), and EC30 (ACAH00000000); *E. faecalis* ARO1/DG (ACAK00000000), ATCC 4200 (ACAG00000000), CH188 (ACAV00000000), D6 (ACAT00000000), DS5 (ACAI00000000), E1Sol (ACAQ00000000); Fly1 (ACAR00000000), HIP11704 (ACAN00000000), JH1 (ACAP00000000), Merz96 (ACAM00000000), T1 (ACAD00000000), T2 (ACAE00000000), T3 (ACAF00000000), T8 (ACOC00000000), T11 (ACAU00000000); and X98 (ACAW00000000); *E. faecium* 1,141,733 (ACAZ00000000), 1,230,933 (ACAS00000000), 1,231,408 (ACBB00000000), 1,231,410 (ACBA00000000), 1,231,501 (ACAY00000000), 1,231,502 (ACAX00000000), Com12 (ACBC00000000), and Com15 (ACBD00000000); and *E. gallinarum* EG2 (ACAJ00000000).

This project has been funded in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under contract no. HHSN266200400001C, and by grants R01AI072360 and P01AI083214 to M.S.G. and EY007145 to K.L.P.

REFERENCES

1. **Aarestrup, F. M., P. Butaye, and W. Witte.** 2002. Nonhuman reservoirs of enterococci, p. 55–99. *In* M. S. Gilmore (ed.), The enterococci: pathogenesis, molecular biology, and antibiotic resistance. ASM Press, Washington, DC.
2. **Benno, Y., A. Izumi-Kurotani, and M. Yamashita.** 1992. Isolation and identification of intestinal bacteria from Japanese tree frog (*Hlya japonica*) with the special reference to anaerobic bacteria. J. Vet. Med. Sci. **54:**699–702.
3. **Bourgogne, A., D. A. Garsin, X. Qin, K. V. Singh, J. Sillanpaa, S. Yerrapragada, Y. Ding, S. Dugan-Rocha, C. Buhay, H. Shen, G. Chen, G. Williams, D. Muzny, A. Maadani, K. A. Fox, J. Gioia, L. Chen, Y. Shang, C. A. Arias, S. R. Nallapareddy, M. Zhao, V. P. Prakash, S. Chowdhury, H. Jiang, R. A. Gibbs, B. E. Murray, S. K. Highlander, and G. M. Weinstock.** 2008. Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. Genome Biol. **9:**R110.
4. **Cartwright, C. P., F. Stock, G. A. Fahle, and V. J. Gill.** 1995. Comparison of pigment production and motility tests with PCR for reliable identification of intrinsically vancomycin-resistant enterococci. J. Clin. Microbiol. **33:**1931–1933.
5. **Cox, C. R., and M. S. Gilmore.** 2007. Native microbial colonization of *Drosophila melanogaster* and its use as a model of *Enterococcus faecalis* pathogenesis. Infect. Immun. **75:**1565–1576.
6. **Domann, E., T. Hain, R. Ghai, A. Billion, C. Kuenne, K. Zimmermann, and T. Chakraborty.** 2007. Comparative genomic analysis for the presence of potential enterococcal virulence factors in the probiotic *Enterococcus faecalis* strain Symbioflor 1. Int. J. Med. Microbiol. **297:**533–539.
7. **Flannagan, S. E., J. W. Chow, S. M. Donabedian, W. J. Brown, M. B. Perri, M. J. Zervos, Y. Ozawa, and D. B. Clewell.** 2003. Plasmid content of a vancomycin-resistant *Enterococcus faecalis* isolate from a patient also colonized by *Staphylococcus aureus* with a VanA phenotype. Antimicrob. Agents Chemother. **47:**3954–3959.
8. **Gardner, P., D. H. Smith, H. Beer, and R. C. Moellering, Jr.** 1969. Recovery of resistance (R) factors from a drug-free community. Lancet **ii:**774–776.
9. **Hidron, A. I., J. R. Edwards, J. Patel, T. C. Horan, D. M. Sievert, D. A. Pollock, and S. K. Fridkin.** 2008. NHSN annual update: antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006–2007. Infect. Control Hosp. Epidemiol. **29:**996–1011.
10. **Jett, B. D., M. M. Huycke, and M. S. Gilmore.** 1994. Virulence of enterococci. Clin. Microbiol. Rev. **7:**462–478.
11. **Leblanc, D. J.** 2006. *Enterococcus*, p. 175–204. *In* M. Dworkin (ed.), The prokaryotes, 3rd ed., vol. 4. Springer, New York, NY.
12. **McBride, S. M., V. A. Fischetti, D. J. Leblanc, R. C. Moellering, Jr., and M. S. Gilmore.** 2007. Genetic diversity among *Enterococcus faecalis*. PLoS One **2:**e582.
13. **Mohamed, J. A., W. Huang, S. R. Nallapareddy, F. Teng, and B. E. Murray.** 2004. Influence of origin of isolates, especially endocarditis isolates, and various genes on biofilm formation by *Enterococcus faecalis*. Infect. Immun. **72:**3658–3663.
14. **Paulsen, I. T., L. Banerjei, G. S. Myers, K. E. Nelson, R. Seshadri, T. D. Read, D. E. Fouts, J. A. Eisen, S. R. Gill, J. F. Heidelberg, H. Tettelin, R. J. Dodson, L. Umayam, L. Brinkac, M. Beanan, S. Daugherty, R. T. DeBoy, S. Durkin, J. Kolonay, R. Madupu, W. Nelson, J. Vamathevan, B. Tran, J. Upton, T. Hansen, J. Shetty, H. Khouri, T. Utterback, D. Radune, K. A. Ketchum, B. A. Dougherty, and C. M. Fraser.** 2003. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. Science **299:**2071–2074.
15. **Pillar, C. M., and M. S. Gilmore.** 2004. Enterococcal virulence—pathogenicity island of *E. faecalis*. Front. Biosci. **9:**2335–2346.
16. **Tannock, G. W., and G. Cook.** 2002. Enterococci as members of the intestinal microflora of humans, p. 101–132. *In* M. S. Gilmore (ed.), The enterococci: pathogenesis, molecular biology, and antibiotic resistance. ASM Press, Washington, DC.
17. **Weigel, L. M., D. B. Clewell, S. R. Gill, N. C. Clark, L. K. McDougal, S. E. Flannagan, J. F. Kolonay, J. Shetty, G. E. Killgore, and F. C. Tenover.** 2003. Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. Science **302:**1569–1571.